

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4366011号
(P4366011)

(45) 発行日 平成21年11月18日 (2009.11.18)

(24) 登録日 平成21年8月28日 (2009.8.28)

(51) Int.Cl.

F I

H04N	1/46	(2006.01)	H04N	1/40	103C
G06K	9/20	(2006.01)	G06K	9/20	340L
G06T	7/00	(2006.01)	G06T	7/00	100C
G06T	7/40	(2006.01)	G06T	7/40	100A
G09G	5/00	(2006.01)	G09G	5/00	530M

請求項の数 9 (全 20 頁) 最終頁に続く

(21) 出願番号 特願2000-388887 (P2000-388887)
 (22) 出願日 平成12年12月21日 (2000.12.21)
 (65) 公開番号 特開2002-190957 (P2002-190957A)
 (43) 公開日 平成14年7月5日 (2002.7.5)
 審査請求日 平成19年12月21日 (2007.12.21)

(73) 特許権者 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100076428
 弁理士 大塚 康德
 (74) 代理人 100112508
 弁理士 高柳 司郎
 (74) 代理人 100115071
 弁理士 大塚 康弘
 (74) 代理人 100116894
 弁理士 木村 秀二
 (74) 代理人 100130409
 弁理士 下山 治
 (74) 代理人 100134175
 弁理士 永川 行光

最終頁に続く

(54) 【発明の名称】 文書処理装置及び方法

(57) 【特許請求の範囲】

【請求項 1】

輝度画像を2値化するための閾値を複数決定する第1決定手段と、

前記第1決定手段で決定された複数の閾値それぞれを用いて前記輝度画像を2値化することにより、複数の2値画像を取得する取得手段と、

前記取得手段で取得した複数の2値画像それぞれに含まれる黒画素の連結成分の領域に基づいて、下地の輝度が異なる各領域の位置と大きさとを示す領域情報を生成する生成手段と、

前記輝度画像における前記領域情報で示される各領域について、輝度値の反転を行うか否かを決定する第2決定手段と、

前記輝度画像における前記領域情報で示される各領域について、当該各領域を2値化するための閾値をそれぞれ設定する設定手段と、

前記輝度画像における前記領域情報で示される各領域について、前記第2決定手段で反転を行うと決定された領域に関しては、輝度値を反転した前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で2値化することにより2値部分画像を取得し、一方、前記第2決定手段で反転を行わないと決定された領域に関しては、輝度値を反転していない前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で2値化することにより2値部分画像を取得する2値部分画像取得手段と、

前記輝度画像の全面を所定閾値で2値化して得られる2値画像における対応する領域に

10

20

、前記 2 値部分画像取得手段で取得した各領域の 2 値部分画像を上書きすることにより、前記輝度画像に関する第 2 の 2 値画像を得る 2 値化手段と、

前記 2 値化手段で得られた第 2 の 2 値画像から文書要素を抽出し、前記生成手段で生成された領域情報で示される各領域を枠の要素とし、前記文書要素と枠の要素とに基づいて木構造を得て、当該木構造に基づいて前記輝度画像の領域分割を行う分割処理手段とを備え、

前記分割処理手段で得る木構造は、前記抽出した文書要素のうち前記領域情報で示される各領域に含まれる文書要素を当該各領域に対応する枠の要素の子孫にした木構造であることを特徴とする文書処理装置。

【請求項 2】

前記生成手段は、前記複数の 2 値画像それぞれに含まれる黒画素の連結成分の領域が所定サイズより大きい領域を、前記下地の輝度が異なる領域と判断して、当該下地の輝度が異なる領域の位置と大きさとを表す領域情報を生成する

ことを特徴とする請求項 1 に記載の文書処理装置。

【請求項 3】

前記設定手段は、前記輝度画像における前記領域情報で示される各領域のヒストグラムに基づいて、当該各領域を 2 値化するための閾値をそれぞれ設定する

ことを特徴とする請求項 1 または 2 に記載の文書処理装置。

【請求項 4】

前記第 2 決定手段は、前記輝度画像における前記領域情報で示される各領域のヒストグラムに基づいて、当該各領域について輝度値の反転を行うか否かを決定する

ことを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の文書処理装置。

【請求項 5】

前記分割処理手段は、

前記 2 値化手段で得られた第 2 の 2 値画像から前記文書要素を抽出して木構造を形成する形成手段と、

前記形成手段で形成した木構造の中に前記領域情報で示される各領域を枠の要素として挿入し、前記抽出された文書要素のうち前記領域情報で示される各文書要素を当該各領域に対応する枠の要素の子孫に移動することにより、前記木構造を変更する変更手段とを備え、

前記変更手段で変更された木構造に基づいて領域分割を行う

ことを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の文書処理装置。

【請求項 6】

カラー画像を前記輝度画像に変換する変換手段を更に備えることを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の文書処理装置。

【請求項 7】

前記第 1 決定手段は、前記輝度画像のヒストグラムをとり、当該ヒストグラムから前記複数の閾値を決定することを特徴とする請求項 1 乃至 6 のいずれか 1 項に記載の文書処理装置。

【請求項 8】

第 1 決定手段が、輝度画像を 2 値化するための閾値を複数決定する第 1 決定工程と、

取得手段が、前記第 1 決定工程で決定された複数の閾値それぞれを用いて前記輝度画像を 2 値化することにより、複数の 2 値画像を取得する取得工程と、

生成手段が、前記取得工程で取得した複数の 2 値画像それぞれに含まれる黒画素の連結成分の領域に基づいて、下地の輝度が異なる各領域の位置と大きさとを示す領域情報を生成する生成工程と、

第 2 決定手段が、前記輝度画像における前記領域情報で示される各領域について、輝度値の反転を行うか否かを決定する第 2 決定工程と、

設定手段が、前記輝度画像における前記領域情報で示される各領域について、当該各領域を 2 値化するための閾値をそれぞれ設定する設定工程と、

10

20

30

40

50

2 値部分画像取得手段が、前記輝度画像における前記領域情報で示される各領域について、前記第 2 決定工程で反転を行うと決定された領域に関しては、輝度値を反転した前記輝度画像内の対応する領域を、前記設定工程で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得し、一方、前記第 2 決定工程で反転を行わないと決定された領域に関しては、輝度値を反転していない前記輝度画像内の対応する領域を、前記設定工程で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得する 2 値部分画像取得工程と、

2 値化手段が、前記輝度画像の全面を所定閾値で 2 値化して得られる 2 値画像における対応する領域に、前記 2 値部分画像取得工程で取得した各領域の 2 値部分画像を上書きすることにより、前記輝度画像に関する第 2 の 2 値画像を得る 2 値化工程と、

10

分割処理手段が、前記 2 値化工程で得られた第 2 の 2 値画像から文書要素を抽出し、前記生成工程で生成された領域情報で示される各領域を枠の要素とし、前記文書要素と枠の要素とに基づいて木構造を得て、当該木構造に基づいて前記輝度画像の領域分割を行う分割処理工程とを有し、

前記分割処理工程で得る木構造は、前記抽出した文書要素のうち前記領域情報で示される各領域に含まれる文書要素を当該各領域に対応する枠の要素の子孫にした木構造であることを特徴とする文書処理方法。

【請求項 9】

コンピュータを、

輝度画像を 2 値化するための閾値を複数決定する第 1 決定手段、

20

前記第 1 決定手段で決定された複数の閾値それぞれを用いて前記輝度画像を 2 値化することにより、複数の 2 値画像を取得する取得手段と

前記取得手段で取得した複数の 2 値画像それぞれに含まれる黒画素の連結成分の領域に基づいて、下地の輝度が異なる各領域の位置と大きさを示す領域情報を生成する生成手段、

前記輝度画像における前記領域情報で示される各領域について、輝度値の反転を行うか否かを決定する第 2 決定手段、

前記輝度画像における前記領域情報で示される各領域について、当該各領域を 2 値化するための閾値をそれぞれ設定する設定手段、

前記輝度画像における前記領域情報で示される各領域について、前記第 2 決定手段で反転を行うと決定された領域に関しては、輝度値を反転した前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得し、一方、前記第 2 決定手段で反転を行わないと決定された領域に関しては、輝度値を反転していない前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得する 2 値部分画像取得手段

30

、前記輝度画像の全面を所定閾値で 2 値化して得られる 2 値画像における対応する領域に、前記 2 値部分画像取得手段で取得した各領域の 2 値部分画像を上書きすることにより、前記輝度画像に関する第 2 の 2 値画像を得る 2 値化手段、

前記 2 値化手段で得られた第 2 の 2 値画像から文書要素を抽出し、前記生成手段で生成された領域情報で示される各領域を枠の要素とし、前記文書要素と枠の要素とに基づいて木構造を得て、ここで、前記木構造は、前記抽出した文書要素のうち前記領域情報で示される各領域に含まれる文書要素を当該各領域に対応する枠の要素の子孫にした木構造であり、当該木構造に基づいて前記輝度画像の領域分割を行う分割処理手段、

40

として機能させるためのプログラムを格納した、コンピュータ読取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、文書の電子化処理を行うための文書処理装置及び方法に関し、特に文書の電子化処理に際して実行される像域分離処理に関する。

50

【 0 0 0 2 】

【従来の技術】

近年、情報の電子化が進み、文書を紙ではなく電子化して保存、あるいは送信するという需要が高まっている。特に、蓄積メディアの低価格化、および通信帯域の増大により、電子化の対象となる文書は白黒 2 値のものから、フルカラーの文書へと広がりつつある。

【 0 0 0 3 】

ここでいう文書の電子化とは、単に紙文書をスキャナなどにより光電変換し、画像データ化することのみにとどまらず、記載されている内容を認識して、文書を構成するテキスト、記号、図、写真、表などそれぞれ性質の異なる領域に分割し、文字部は文字コード情報、図はベクトルデータ、写真は画像データ、表は構造データ、といったように各々最も適した形態でデータ化する処理のことを指す。

10

【 0 0 0 4 】

そのような文書の電子化処理の第 1 段階として行われるのが、1 頁の文書画像に書かれた内容を分析し、文字や図、写真、表など異なる性質の部分要素に分割する処理、すなわち領域分割処理である。図 2 5 は領域分割の一例を示す図である。

【 0 0 0 5 】

このような領域分割処理の実現例としては、米国特許第 5 6 8 0 4 7 8 号の“Method and Apparatus for character recognition”(Shin-Ywangら / Canon K.K)などが挙げられる。この例では、文書画像中の黒画素の 8 連結輪郭塊、白画素の 4 連結輪郭塊の集合を抽出し、その形状、大きさ、集合状態などから、文字領域、絵や図、表、枠、線といった文書に特徴的な領域を抽出している。図 2 5 の例では、文字領域(ブロック 1、3、4、6)、絵や図領域(ブロック 2)、表領域(ブロック 5)、枠、線(7)といった文書に特徴的な領域を抽出している。

20

【 0 0 0 6 】

ここで、黒画素の 8 連結輪郭塊(以下、黒画素塊)とは、図 1 4 のように、ある黒画素から 8 方向のいずれかで連結している黒画素の集合体である。また、白画素の 4 連結輪郭塊(以下、白画素塊)とは、図 1 6 のようにある白画素から 4 方向のいずれかで連結している白画素の集合体を指す。

【 0 0 0 7 】

上述の領域分割処理は、その動作原理より、入力となる文書画像が白黒 2 値であることが前提となる。従って、この技術を利用してカラー文書の領域分割を行うためには、あらかじめ文書画像の 2 値化を行う必要がある。一般にカラー画像の 2 値化は、画素の輝度分布から閾値を求め、画像の各画素を、前出の輝度閾値を境に白又は黒の画素へと変換することで行われる。

30

【 0 0 0 8 】

【発明が解決しようとする課題】

カラー画像の 2 値化のための閾値の求め方には、全画面単一で求める方法と、領域毎に求める方法が考えられる。本出願人による特願平 1 1 - 2 3 8 5 8 1 号にて提案されている 2 値化方法では、入力原稿の内容に応じて、領域毎に動的に最適な閾値を求め、これを用いることで領域毎に最適な 2 値化が行われる。特に低輝度下地上の高輝度文字と、高輝度下地上の低輝度文字が混在するカラー画像から、それらすべてが白地上の黒文字に自動的に変換されるような 2 値化を可能とし、領域分割処理の入力として最適な 2 値画像を得ることが可能である。

40

【 0 0 0 9 】

図 2 4 は先に提案された 2 値化手法によって色付きの下地を含む文書に領域分離を行った様子を説明する図である。図 2 4 において、カラーの文書 2 3 0 1 は、下半分に濃い色付の下地領域があり、その上に薄い色の文字が載っているものとし、それ以外は薄い色の下地に濃い色の文字となっているものとする。このような文書においては、上半分と下半分は意味的に分離しているであろうことが理解できる。

【 0 0 1 0 】

50

文書 2301 のようなカラー文書に対して、前述の 2 値化方式で 2 値化を行うと、図 24 の 2302 で示されるような 2 値画像が生成される。2 値画像 2302 では、下地は除去されてすべて白画素になり、文字はすべて黒画素になっている。このとき、2 値画像 2302 に対し従来通り領域分割処理を行うと、図 24 中 2303 のような結果が得られる、画面の下半分に存在した下地付の領域の情報が欠落しているために、本来 T E X T 1 および T E X T 2 はそれぞれ中央で 2 つに分離すべきなのにもかかわらず、結合してしまっている。

【0011】

つまり、本来カラー画像の持っている下地色による文字領域の範囲指定情報は 2 値化の際に失われてしまう。

【0012】

本発明は上記の問題に鑑みてなされたものであり、色によって表された領域の区別を維持した領域分割を可能にすることを目的とする。

【0013】

【課題を解決するための手段】

上記の目的を達成するための本発明の一態様による文書処理装置は、

輝度画像を 2 値化するための閾値を複数決定する第 1 決定手段と、

前記第 1 決定手段で決定された複数の閾値それぞれを用いて前記輝度画像を 2 値化することにより、複数の 2 値画像を取得する取得手段と、

前記取得手段で取得した複数の 2 値画像それぞれに含まれる黒画素の連結成分の領域に基づいて、下地の輝度が異なる各領域の位置と大きさを示す領域情報を生成する生成手段と、

前記輝度画像における前記領域情報で示される各領域について、輝度値の反転を行うか否かを決定する第 2 決定手段と、

前記輝度画像における前記領域情報で示される各領域について、当該各領域を 2 値化するための閾値をそれぞれ設定する設定手段と、

前記輝度画像における前記領域情報で示される各領域について、前記第 2 決定手段で反転を行うと決定された領域に関しては、輝度値を反転した前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得し、一方、前記第 2 決定手段で反転を行わないと決定された領域に関しては、輝度値を反転していない前記輝度画像内の対応する領域を、前記設定手段で当該領域に対して設定された閾値で 2 値化することにより 2 値部分画像を取得する 2 値部分画像取得手段と、

前記輝度画像の全面を所定閾値で 2 値化して得られる 2 値画像における対応する領域に、前記 2 値部分画像取得手段で取得した各領域の 2 値部分画像を上書きすることにより、前記輝度画像に関する第 2 の 2 値画像を得る 2 値化手段と、

前記 2 値化手段で得られた第 2 の 2 値画像から文書要素を抽出し、前記生成手段で生成された領域情報で示される各領域を枠の要素とし、前記文書要素と枠の要素とに基づいて木構造を得て、当該木構造に基づいて前記輝度画像の領域分割を行う分割処理手段とを備え、

前記分割処理手段で得る木構造は、前記抽出した文書要素のうち前記領域情報で示される各領域に含まれる文書要素を当該各領域に対応する枠の要素の子孫にした木構造である。

【0014】

また、上記の目的を達成するための本発明の文書処理方法は、

第 1 決定手段が、輝度画像を 2 値化するための閾値を複数決定する第 1 決定工程と、

取得手段が、前記第 1 決定工程で決定された複数の閾値それぞれを用いて前記輝度画像を 2 値化することにより、複数の 2 値画像を取得する取得工程と、

生成手段が、前記取得工程で取得した複数の 2 値画像それぞれに含まれる黒画素の連結成分の領域に基づいて、下地の輝度が異なる各領域の位置と大きさを示す領域情報を生

10

20

30

40

50

成する生成工程と、

第2決定手段が、前記輝度画像における前記領域情報で示される各領域について、輝度値の反転を行うか否かを決定する第2決定工程と、

設定手段が、前記輝度画像における前記領域情報で示される各領域について、当該各領域を2値化するための閾値をそれぞれ設定する設定工程と、

2値部分画像取得手段が、前記輝度画像における前記領域情報で示される各領域について、前記第2決定工程で反転を行うと決定された領域に関しては、輝度値を反転した前記輝度画像内の対応する領域を、前記設定工程で当該領域に対して設定された閾値で2値化することにより2値部分画像を取得し、一方、前記第2決定工程で反転を行わないと決定された領域に関しては、輝度値を反転していない前記輝度画像内の対応する領域を、前記設定工程で当該領域に対して設定された閾値で2値化することにより2値部分画像を取得する2値部分画像取得工程と、

10

2値化手段が、前記輝度画像の全面を所定閾値で2値化して得られる2値画像における対応する領域に、前記2値部分画像取得工程で取得した各領域の2値部分画像を上書きすることにより、前記輝度画像に関する第2の2値画像を得る2値化工程と、

分割処理手段が、前記2値化工程で得られた第2の2値画像から文書要素を抽出し、前記生成工程で生成された領域情報で示される各領域を枠の要素とし、前記文書要素と枠の要素とに基づいて木構造を得て、当該木構造に基づいて前記輝度画像の領域分割を行う分割処理工程とを有し、

前記分割処理工程で得る木構造は、前記抽出した文書要素のうち前記領域情報で示される各領域に含まれる文書要素を当該各領域に対応する枠の要素の子孫にした木構造である

20

【0015】

【発明の実施の形態】

以下、添付の図面を参照して本発明の好適な実施形態を説明する。

【0016】

図1は本実施形態による文書処理装置の装置概観を示す図である。図1において、101はコンピュータ装置であり、添付のフローチャートを参照して以下で説明する処理を実現するためのプログラムを含む、文書の電子化処理プログラムを実行する。コンピュータ装置101は、ユーザに状況や画像を表示するためのディスプレイ装置102、およびユーザの操作を受け付けるキーボードやマウス等のポインティングデバイスを含んで構成される入力装置103が付随する。ディスプレイ102表示デバイスとしては、CRTやLCD等が用いられる。104はスキャナ装置であり、文書画像を光学的に読み取り、電子化し、得られた画像データをコンピュータ装置101に送る。本実施形態では、カラーキャナを用いる。

30

【0017】

図2は本実施形態による文書処理装置の構成を示すブロック図である。図1において、201はCPUであり、ROM202或いはRAM203に格納された制御プログラムを実行することにより、後述の電子化処理を含む各種機能を実現する。202はROMであり、CPU201によって実行される各種制御プログラムやデータが格納される。203はRAMであり、CPU201によって実行される各種制御プログラムを格納したり、CPU201が各種処理を実行するのに必要な作業領域を提供する。204は外部記憶装置であり、添付のフローチャートを参照して説明する処理をCPU101によって実現するための制御プログラムや、画像入力装置104を読み取って得られた文書画像データ等を格納する。205はコンピュータバスであり、上述の各構成を接続する。

40

【0018】

図3は本実施形態の文書処理装置による文書の電子化処理の概要を示す図である。本実施形態による電子化処理の流れは、まず入力部301において電子化の対象とするカラー文書をスキャナ104を用いて読み込み、文書を画像データとして外部記憶装置204に格納する。次に、2値化部302において、後段の領域分割処理のために外部記憶装置20

50

4に格納された文書画像データに対して2値化処理を施す。領域分割部303では、2値化部302で得られた2値画像から、文字、図、表、枠、線などの要素を抽出し、各領域に分割する。電子化文書作成部304は、分割された各要素毎の属性に基づいて文字認識データや用いたり表構造データを用いたりして電子化文書を作成する。出力部305は生成された電子化文書を外部記憶装置204に格納する。なお、出力部305における出力の形態は外部記憶装置204への格納に限られるものではなく、ディスプレイ102へ表示出力したり、不図示のネットワークインターフェースを介してネットワーク上の他の装置へ出力したり、不図示のプリンタへ出力したりすることも可能である。

【0019】

以下、図3に示した2値化部302の動作について図面参照して説明する。図4は本実施形態による2値化処理を説明するフローチャートである。また、図5は本実施形態の説明に用いるカラー文書画像の例を示す図である。なお、図5に示すカラー文書画像は、3つの下地色（下地A501、下地B502、下地C503）を含み、下地A501の部分が白、下地B502の部分が黄色、下地C503の部分が青であるとする。一方、文字色は、文字列A504、文字列B505が共に黒、文字列C506は白である。

【0020】

まず、ステップS401にて処理対象のカラー文書画像を輝度画像に変換する。ここでは一般的に元画像のカラー形式をRGB形式、輝度画像は1画素あたり0～255のグレイスケール形式とし、各画素の輝度Yは元画像の画素値R、G、Bに対し、 $Y = 0.299R + 0.587G + 0.114B$ によって求めることにするが、他の形式/変換式を用いてもよいことはもちろんである。これにより図5の下地色A501、下地色B502、下地色C503の各部分の輝度はそれぞれ255、200、40（実際は紙面状態や光電変換時のばらつきによって、同じ下地色であっても位置によって輝度値がばらつく。したがって、図6のようなヒストグラム曲線の山は、ある程度の幅を有することになる）へと変換される。

【0021】

次に、ステップS402にて、ステップS401で得られた輝度画像データのヒストグラムをとる。図6は図5に示したカラー画像から得られる輝度値のヒストグラムを示す図である。ヒストグラムが得られたならば、ステップS403において、そのヒストグラムから複数の閾値を決定する。例えば図6に示したヒストグラムからは、 t_1 、 t_2 の2個の閾値が抽出される。

【0022】

この複数の閾値 t_1 、 t_2 は例えば以下のような手順で決定できる。ヒストグラムの輝度値255（白）から0（黒）まで参照する。参照した輝度値から閾値を選択する条件は、例えば、以下の各条件を設定する。

【0023】

条件1：当該ヒストグラム曲線において、現在の参照点と、その参照点から横軸の輝度値の大きい方向に例えば10画素分だけさかのぼった点との間に含まれる合計度数（面積）が、所定の第1の値より大きい。

【0024】

条件2：当該ヒストグラム曲線における縦軸の度数の急激な減少が発生（例えば「所定の第1の傾きより大」が1回、または「所定の第2の傾きより大」が連続して2回発生）。

【0025】

条件3：当該ヒストグラム曲線の緩い減少、または上昇が発生（所定の第3の傾きより小）。

【0026】

本実施形態において、図6に示す t_1 、 t_2 の各点は、上記の各条件のうち、「条件1」が満足された後に「条件2」が満足され、その後更に「条件3」が満足された結果として検出した。尚、参照点は、当該ヒストグラム曲線を所定の間隔で参照するようにしてもよい。また、本実施形態では以上のような条件を用いたが、閾値の決定はこれに限るもので

10

20

30

40

50

はない。例えば更に以下のような条件 4 ~ 6 を同時に満たす場合を閾値として選択しても良い。

【 0 0 2 7 】

条件 4 : 当該ヒストグラム曲線において、現在の参照点と、その参照点から横軸の輝度値の大きい方向に例えば 4 0 画素さかのぼった点との間に含まれる合計度数 (面積) が、所定の第 2 の値より大きい。

【 0 0 2 8 】

条件 5 : 当該ヒストグラム曲線上の現在の参照点における縦軸の度数が、所定の第 3 の値より小さい。

【 0 0 2 9 】

条件 6 : 当該ヒストグラム曲線において、現在の参照点と、その参照点から横軸の輝度値の大きい方向に例えば 2 0 画素さかのぼった点におけるヒストグラム値が、所定の第 4 の値より大きい。

【 0 0 3 0 】

ステップ S 4 0 4 にて、各閾値から複数の 2 値画像を作成する。これらはテンポラリの 2 値画像である。本例では、閾値 t_1 で 2 値化することにより、図 7 の (a) に示す 2 値画像 7 0 1 が、閾値 t_2 で 2 値化することにより図 7 の (b) に示す 2 値画像 7 0 2 が生成される。ここで、領域 7 0 3 及び 7 0 4 はそれぞれ図 5 に示す下地 B 5 0 2、C 5 0 3 の領域である。閾値 t_1 では下地 B 5 0 2 と下地 C 5 0 3 の部分も黒となり、閾値 t_2 では、下地 5 0 2 の部分は白となり、下地 C 5 0 3 の部分は黒となる。

【 0 0 3 1 】

次に、ステップ S 4 0 5 にて、ステップ S 4 0 4 で得られた各 2 値画像中の黒領域を抽出し、これを下地色の異なる領域を示す領域情報として記録する。これは、2 値画像上の黒画素の連結成分を追跡して、一定の大きさ以上となる四角形の領域を抽出する処理である。図 7 の例では、2 値画像 7 0 1 より領域 7 0 3 が、2 値画像 7 0 2 より領域 7 0 4 がそれぞれ抽出される。

【 0 0 3 2 】

次に、ステップ S 4 0 6 にて、ステップ S 4 0 5 で抽出した各領域内の解析を行う。ここでは、ステップ S 4 0 1 で得られた輝度画像より、ステップ S 4 0 5 で抽出された各領域の範囲を切り出し、各範囲毎に輝度値のヒストグラムをとり、このヒストグラムから各領域について輝度値の反転が必要か否か判断し、必要な場合は輝度値を反転する。更に、このヒストグラムから各領域内に最適な 2 値化閾値を、図 6 と同様にして求める (ただし、対象とする領域の大きさが小さくなるので判断に用いる所定の値は異なる)。求められた 2 値化閾値及び反転情報は、各領域情報に付随するように出力される。尚、反転が必要か否かは、例えば以下に示す計算式により、ヒストグラムの平均値 (average) とそのスキュー (skew) とを求めて判断できる。

【 0 0 3 3 】

$$\begin{aligned} (\text{average}) &= \sum ip(i), \\ \sigma^2 &= \sum (i - \text{av})^2 p(i), \\ (\text{skew}) &= 1 / \sigma^3 \sum (i - \text{av})^3 p(i), \end{aligned}$$

但し、 $\sum ip(i)$ は、 $ip(i)$ の $i = 0$ から $i = 255$ までの場合の総和、 $p(i)$ は確率密度、 av は平均値を表わす。

【 0 0 3 4 】

このスキューの絶対値がしきい値より大きいときには、その領域は文字を含むと判断し、更に文字を含むと判断されて且つスキューが正のときには、領域の画像を反転すべきと判断して反転フラグが yes になる。

【 0 0 3 5 】

本例では、領域 7 0 3 より図 8 に示されるヒストグラムが求められて反転必要なしと判断されて該ヒストグラムから閾値 t_{21} が求められる。さらに、領域 7 0 4 より図 9 に示されるヒストグラムが求められて反転必要ありと判断され、図 9 のヒストグラムを反転して

閾値 t_{22} が求められる（尚、図 9 は反転前のヒストグラムを表し、閾値 t_{22} は反転前のヒストグラムの対応する値として示している）。そして、図 10 に示すように、それぞれの領域毎に、座標値と 2 値化閾値、反転情報を 1 組とした領域情報が記憶される。

【0036】

本例では、領域 703 に対する領域情報 1001 は、座標値として矩形の左端の x 座標 l_1 、右端の x 座標 r_1 、上端の t_1 、下端の b_1 を、2 値化閾値として図 8 のヒストグラムから求めた閾値 t_{21} を、反転情報として反転無し (no) を有する。同様に、領域 704 に対する領域情報 1002 は、座標値として矩形の左端の x 座標 l_2 、右端の x 座標 r_2 、上端の t_2 、下端の b_2 を、2 値化閾値として図 9 のヒストグラムから求めた閾値 t_{22} を、反転情報として反転有り (yes) を有する。

10

【0037】

次いで、ステップ S407 にて、ステップ S406 で得られた領域情報を基に、ステップ S401 で得た輝度画像から最終的な 2 値画像を生成する。本例によれば、

(1) 全面を閾値 t_1 で 2 値化して 2 値画像 701 を得る。

(2) 2 値画像 701 上の領域 703 の部分の輝度画像について、領域情報 1001 の内容にしたがって処理（この場合閾値 (t_{21}) で 2 値化）を行い、得られた画像を領域 703 の部分に上書きする。この結果、2 値画像 702 が得られる。

(3) 領域 704 については、対応する領域情報 1002 が反転有りを示すので、領域 704 の部分の輝度画像について反転処理を行い、閾値 t_{22} を用いて 2 値化して得られた画像を領域 704 の部分に上書きする。この結果、図 5 に示したカラー画像から図 11 に示すような最終 2 値画像が得られることになる。

20

【0038】

次に、以上のようにして 2 値化部 302 によって得られた 2 値画像と領域情報を用いて、領域分割部 303 が領域分割を行う。以下、領域分割部 303 の処理について説明する。

【0039】

図 12 は本実施形態の領域分割処理を説明するフローチャートである。領域分割部 303 では、ステップ S1201 にて、2 値化部 302 により得られた 2 値化された文書画像から文書要素が抽出され、それらの木構造表現が作られる。このステップ S1201 の処理の詳細について図 13 のフローチャートを用いて説明する。図 13 は図 12 のステップ S1201 の要素抽出及び木構造化の処理を説明するフローチャートである。

30

【0040】

ステップ S1301 では、2 値化画像からすべての黒画素塊を抽出する。なお、黒画素塊とは、上述したように黒画素の 8 連結輪郭塊であり、図 14 に示すように、縦横斜めに接触した画素で作られた輪郭を持つ黒画素の集合のことである。続くステップ S1302 では、抽出した黒画素塊が、予め予想される最大文字高さおよび幅（予め実験的に求めた値）に対し定められた閾値以下の大きさを有するかどうかを判定し、閾値以下である場合はステップ S1308 に進み、当該黒画素塊を文字要素と判定する。これを“CHAR”と呼ぶ。

【0041】

ステップ S1303 では、抽出した黒画素が一定比率以上で縦長あるいは横長であるかどうかを判断する。当該黒画素塊が一定比率以上で縦長あるいは横長であった場合は、ステップ S1309 において“LINE”と判定する。また、ステップ S1304 では、抽出した黒画素塊中の黒画素のなす輪郭に注目し、その形状が細い斜めの線状であった場合は、ステップ S1309 へ進み、当該黒画素塊を“LINE”と判定する。

40

【0042】

ステップ S1305 では、黒画素塊の輪郭形状が四角形かどうかを調べる。図 15 は、(a) 黒画素塊の輪郭が四角形の場合と、(b) 黒画素塊の輪郭が非四角形である場合の例を示す図である。ステップ S1305 において、黒画素塊の輪郭形状が四角形でなければ、ステップ S1312 へ進み、当該黒画素塊を“PICTURE”と判定する。

【0043】

50

一方、黒画素塊が四角形ならばステップS 1 3 0 6へ進む。ステップS 1 3 0 6では、黒画素塊の内部に存在する白画素の4連結輪郭塊を抽出する。白画素の4連結輪郭塊とは、図16のように、縦横のみに接触した画素で作られた輪郭を持つ白画素の集合のことである。以降この集合を白画素塊と呼ぶ。

【0044】

ステップS 1 3 0 7では、ステップS 1 3 0 6で黒画素塊から抽出された白画素塊の形状がすべて四角形であり、かつ黒画素塊内を所定の間隔で隙間なく埋めているかどうかを判定する。この判定の結果がYESであった場合は、ステップS 1 3 1 1へ進み、当該黒画素塊を“FRAME”と判定する。図17は、枠(FRAME)と図(PICTURE)における内部白画素塊の配置例を示す図である。ステップS 1 3 0 7の判定によれば、(a)、(b)はステップS 1 3 0 7の判定条件を満たすので、ステップS 1 3 1 1で枠(FRAME)であると判定されることになる。

10

【0045】

また、図17の(c)に示した白画素塊の配列は、「黒画素塊から抽出された白画素塊の形状がすべて四角形であり、かつ黒画素塊内を隙間なく埋めている」という条件を満たしておらず、ステップS 1 3 1 2において図(PICTURE)と判定されることになる。結局、ステップS 1 3 1 2では、ステップS 1 3 0 7までの条件のいずれにもあてはまらない黒画素塊を、“PICTURE”としている。

【0046】

さて、本実施形態では、各要素の親となる要素“GROUND”を導入する。画面全体をひとつの“GROUND”とすると、これまで画像から抽出された各要素は、すべてその子供の要素として表現される。そして、“FRAME”と判定された黒画素塊の内部から抽出された白画素塊のひとつひとつをそれぞれ“GROUND”とし、更にこの白画素の内部で、上述したステップS 1 3 0 1～S 1 3 1 2の処理を行って子供となる要素を抽出する(ステップS 1 3 1 3)。なお、“FRAME”の内部で更に“FRAME”が抽出されたときには、これをGROUNDとしてさらに再帰的に処理を行う。

20

【0047】

すべての再帰的内部探索が終了した時点で、画像から抽出された要素は木構造を構成することになる。図18は、文書画像の例と、これをステップS 1 2 0 1の要素抽出、木構造化によって処理して得られる木構造の例を示す図である。図18の(a)に示されるように、文書画像1801は、テキスト列(CHAR)1802、1807、1808と、フレーム(FRAME)1804と、図(PICTURE)1803、1809とを有する。

30

【0048】

この文書画像1801を上述の処理によって木構造化すると、図18の(b)のようになる。GROUND1821は文書画像1801の全体を示し、その要素の一つであるFRAME1824は文書画像1801中のフレーム1804に対応する。更にフレーム1804は2つのフレームに分けられ、木構造上では、それぞれGROUND1825、1826として示されている。

【0049】

以上のようにして、図12のステップS 1 2 0 1で文書画像要素の木構造を得ると、ステップS 1 2 0 2において、ステップS 4 0 6で取得した領域情報を木構造に当てはめる。即ち、2値化部302による2値化処理の際に記録された領域情報を参照して、ステップS 3 0 1で得られた木構造を変更する。このステップS 1 2 0 2の処理について、図19のフローチャートを用いて説明する。

40

【0050】

ステップS 1 9 0 1では、領域情報が存在するかどうかを調べる。領域情報がなければ、そのまま本処理を終了する。領域情報が存在する場合は、ステップS 1 9 0 2へ進み、領域を仮想的に“FRAME”および“GROUND”の組と考えて要素の木構造の適当な箇所に挿入する。すなわち、領域を内包する“GROUND”を親とする位置に、その領域情報が表す矩形に対応する新たな“FRAME”を挿入し、その子として新たな“GROUND”を置く。

50

【 0 0 5 1 】

ステップ S 1 9 0 3 では、ステップ S 1 9 0 2 で挿入した “ FRAME ” と同じ親 (GROUND) を持つ兄弟のうち、当該領域内部に位置する要素をすべて自分の子孫、すなわち新たな “ GROUND ” の子供へと移動する。

【 0 0 5 2 】

図 2 0 は、上述した木構造変更処理の例を説明する図である。文書画像 2 0 0 1 は色付きの下地領域 2 0 0 2 を有する。2 値化部 3 0 2 はこの文書画像 2 0 0 1 を 2 値化して 2 値画像 2 0 1 0 を生成するとともに、領域 2 0 0 2 に対応する領域情報 2 0 2 0 を生成する (図 4)。領域分割部 3 0 3 は、2 値画像 2 0 1 0 について図 1 3 のフローチャートで説明した処理を施すことにより、画像全体を GROUND 2 0 3 1 とする木構造 2 0 3 0 を生成する (S 1 2 0 1)。そして、木構造 2 0 3 0 に、領域情報 2 0 2 0 を当てはめて、木構造を変更し、木構造 2 0 4 0 を得る。

10

【 0 0 5 3 】

より具体的には、GROUND 2 0 3 1 を親として、領域 2 0 0 2 に相当する枠 (FRAME) を挿入して、その子として GROUND 2 0 4 1 を置く。そして、領域 2 0 0 2 に含まれる各要素 (A , B , C , 図形) を GROUND 2 0 4 1 の子として配置することにより、木構造を変更する。

【 0 0 5 4 】

以上のようにして、ステップ S 1 2 0 2 の処理を終えたならば、ステップ S 1 2 0 3 へ進む。ステップ S 1 2 0 3 では、文字要素をグループ化して行および文字領域を作成する。ステップ S 1 2 0 3 の処理について、図 2 1 のフローチャートを用いて説明する。図 2 1 はステップ S 1 2 0 3 における文字領域の作成処理を説明するフローチャートである。

20

【 0 0 5 5 】

ステップ S 2 1 0 1 では、各 “ CHAR ” に対し、それぞれ隣り合う “ CHAR ” との水平距離が閾値内にあるもの同志をグループ化する。このグループを “ TEXTLINE ” と呼ぶ。但し、これらのグループ化は、同じ親を持つ “ CHAR ” 間のみで行われる。

【 0 0 5 6 】

次に、ステップ S 2 1 0 2 において、更にこの “ TEXTLINE ” 同志で、それぞれ隣り合う垂直距離が閾値内にあるもの同志をグループ化する。こうして得られた “ TEXTLINE ” のグループを “ TEXT ” あるいは文字領域と呼ぶ。但し、これらのグループ化は同じ親をもつ “ CHAR ” からなる “ TEXTLINE ” の間のみで行う。

30

【 0 0 5 7 】

以上の処理で、文書画像は文字領域である “ TEXT ”、線の画像部分である “ L LINE ”、図や写真の領域である “ PICTURE ”、表や枠の領域である “ FRAME ” という各要素に分割される。

【 0 0 5 8 】

なお、ステップ S 2 1 0 3 における文字領域作成時のグループ化は、木構造を意識して行われるので、例えば図 2 2 のように、2 値画像上の実線の枠 = “ FRAME 1 ” がある場合、“ FRAME 1 ” 内外それぞれの文字は必ず異なるグループとなる。すなわち、“ TEXT 1 ” と “ TEXT 3 ”、“ TEXT 2 ” と “ TEXT 4 ” のように異なる文字領域にグループ化される。

40

【 0 0 5 9 】

同様に、図 2 3 におけるようなカラー画像の処理の場合、2 値化処理後の画像上では枠となるような情報がなくても、ステップ S 1 2 0 2 における処理により、2 値化処理時に得られた領域情報が領域分割に反映され、色下地上の文字は他と異なる文字領域となり、図 2 2 と同様な正しい領域分割結果が得られる。即ち、2 値化処理部 3 0 2 によって、文書画像 2 3 0 1 の 2 値画像 2 3 1 0 が得られるとともに領域情報 2 3 2 0 が得られる。2 値画像 2 3 1 0 を領域分割する際に、ステップ S 1 2 0 2 の処理により、領域情報 2 3 2 0 が反映されて、2 3 3 0 に示す如く領域分割結果が得られる。

【 0 0 6 0 】

もし、ステップ S 1 2 0 2 の処理を行わないで領域分割処理を行った場合は、下地の情報

50

が反映されることがないので、図 2 4 のように誤った文字領域が得られることになる。

【 0 0 6 1 】

以上説明したように、本実施形態によれば、カラー画像を 2 値化して領域分割処理を行う際に、2 値化処理時に記憶された色付下地の領域情報を用いて領域分割処理の解析内容を変更するので、カラー画像の分割処理に際して 2 値化の際に失われてしまった情報をも正しく反映した文字領域の抽出が可能になり、より高精度な領域分割処理が可能になる。

【 0 0 6 2 】

なお、上記実施形態によれば、2 値化部 3 0 2 が保存する領域情報、および領域分割部 3 0 3 が処理する“FRAME”領域は四角形に限定していたが、これは矩形の任意連接体、あるいは円、楕円などの領域を対象とするようにしてもよい。この場合でも、カラー画像にあって、2 値化の際に失われてしまった情報をも正しく反映した文字領域の抽出が可能になり、より高精度な領域分割処理が可能になる。

【 0 0 6 3 】

なお、本発明は、複数の機器（例えばホストコンピュータ、インタフェイス機器、リーダ、プリンタなど）から構成されるシステムに適用しても、一つの機器からなる装置（例えば、複写機、ファクシミリ装置など）に適用してもよい。

【 0 0 6 4 】

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体（または記録媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（または CPU や MPU）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム（OS）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【 0 0 6 5 】

さらに、記憶媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わる CPU などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【 0 0 6 6 】

【発明の効果】

以上説明したように、本発明によれば、色によって表された領域の区別を維持した領域分割が可能となる。

【図面の簡単な説明】

【図 1】本実施形態による文書処理装置の装置概観を示す図である。

【図 2】本実施形態による文書処理装置の構成を示すブロック図である。

【図 3】本実施形態の文書処理装置による文書の電子化処理の概要を示す図である。

【図 4】本実施形態による 2 値化処理を説明するフローチャートである。

【図 5】本実施形態の説明に用いるカラー文書画像の例を示す図である。

【図 6】図 5 に示したカラー画像から得られる輝度値のヒストグラムを示す図である。

【図 7】（a）は図 5 に示したカラー画像を図 6 に示す閾値 t_1 で 2 値化した画像を示し、（b）は閾値 t_2 で 2 値化した画像を示す図である。

【図 8】図 7 の領域 7 0 3 における輝度値のヒストグラムを示す図である。

【図 9】図 7 の領域 7 0 4 における輝度値のヒストグラムを示す図である。

【図 10】領域情報を示す図である。

【図 1 1】図 5 に示したカラー画像の最終的な 2 値画像を示す図である。

【図 1 2】本実施形態の領域分割処理を説明するフローチャートである。

【図 1 3】図 1 2 のステップ S 1 2 0 1 の要素抽出及び木構造化の処理を説明するフローチャートである。

【図 1 4】黒画素の 8 連結輪郭塊の例を示す図である。

【図 1 5】(a) 黒画素塊の輪郭が四角形の場合と、(b) 黒画素塊の輪郭が非四角形である場合の例を示す図である。

【図 1 6】白画素の 4 連結輪郭塊の例を示す図である。

【図 1 7】枠 (FRAME) と図 (PICTURE) における内部白画素塊の配置例を示す図である。

【図 1 8】文書画像の例と、これをステップ S 1 2 0 1 の要素抽出、木構造化によって処理して得られる木構造の例を示す図である。

【図 1 9】図 1 2 のステップ S 1 2 0 2 による木構造の変更処理を説明するフローチャートである。

【図 2 0】上述した木構造変更処理の例を説明する図である。

【図 2 1】ステップ S 1 2 0 3 における文字領域の作成処理を説明するフローチャートである。

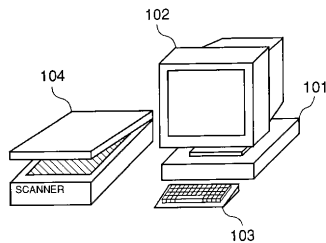
【図 2 2】枠領域を持つ文書の領域分割処理例を示す図である。

【図 2 3】本実施形態により色付き下地領域を持つカラー文書が正しく領域分割される様子を示す図である。

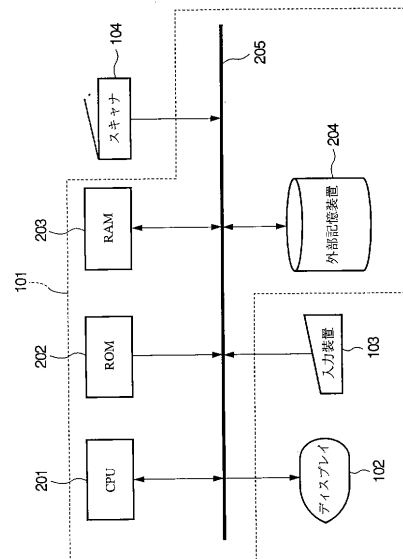
【図 2 4】一般的な手法により色付きの下地領域を持つカラー文書に領域分割を行った様子

【図 2 5】領域分割の一例を示す図である。

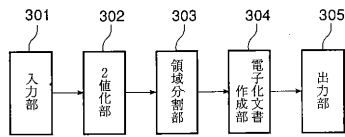
【図 1】



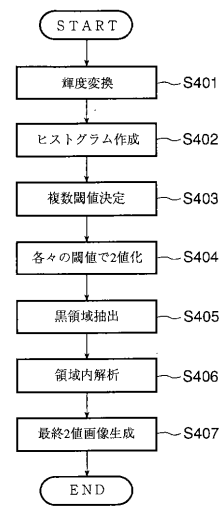
【図 2】



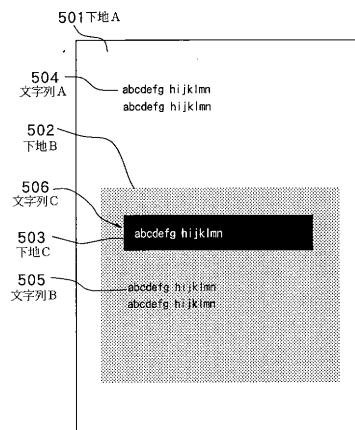
【図 3】



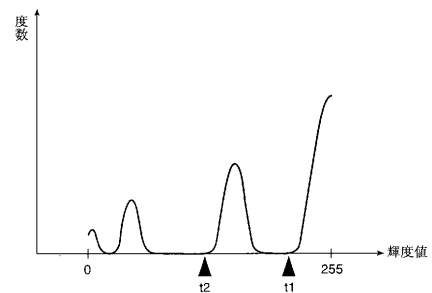
【図 4】



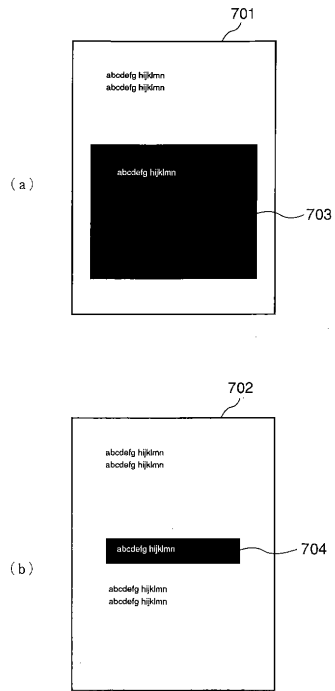
【図 5】



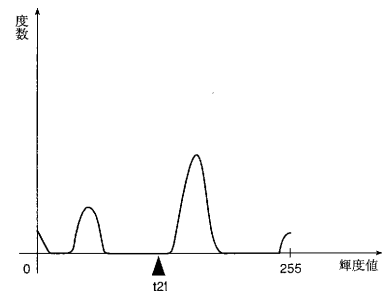
【図 6】



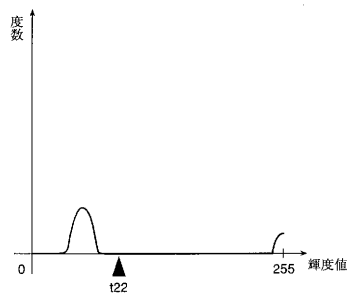
【図 7】



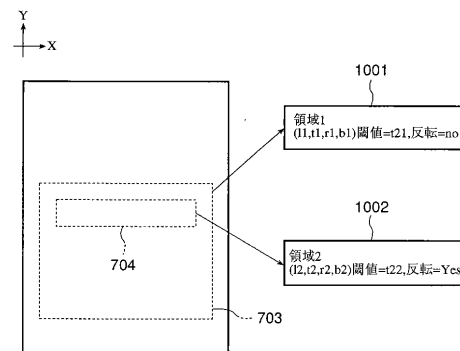
【図 8】



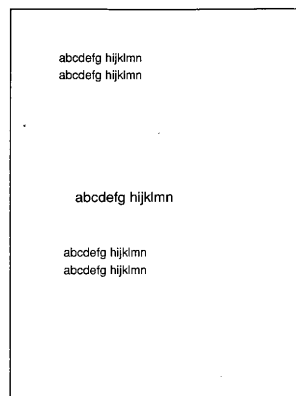
【図 9】



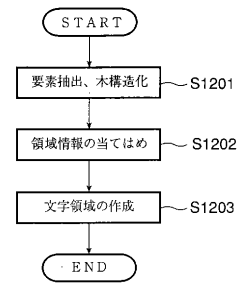
【図 10】



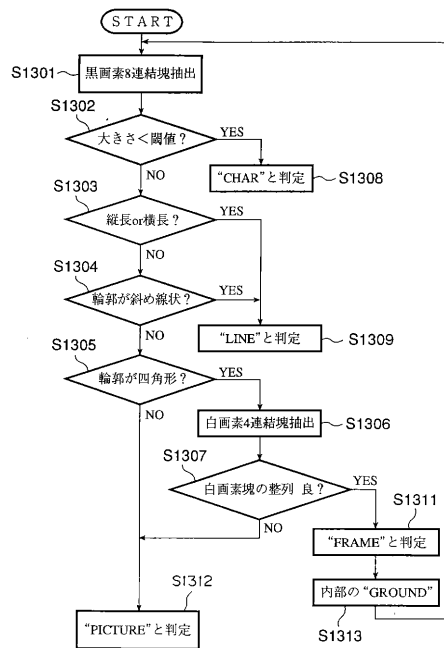
【図 1 1】



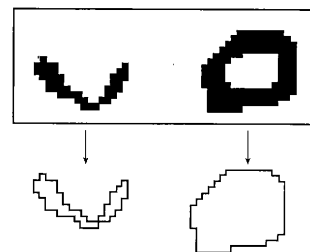
【図 1 2】



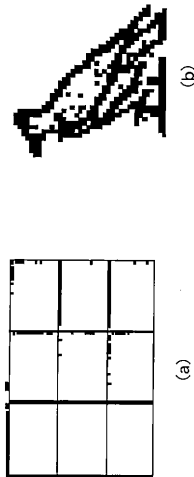
【図 1 3】



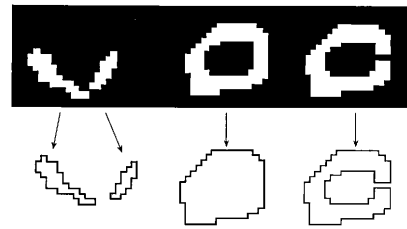
【図 1 4】



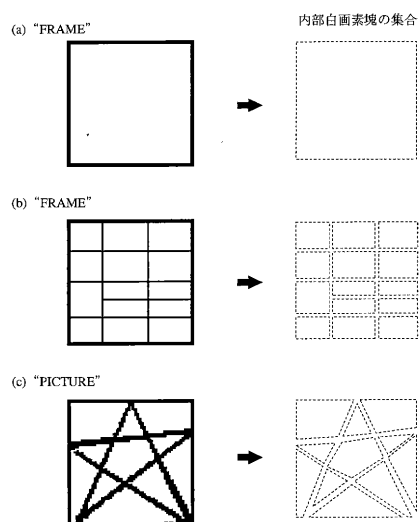
【図 15】



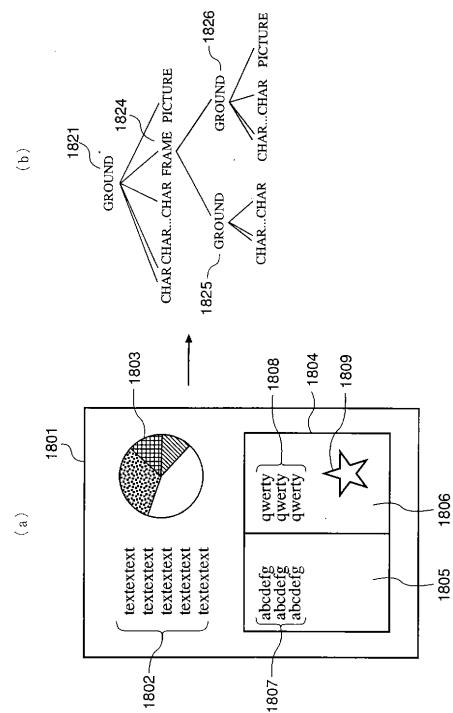
【図 16】



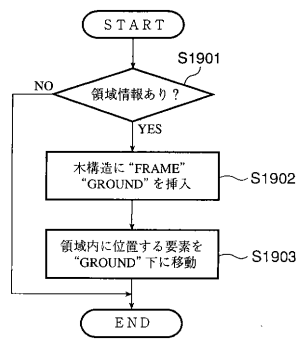
【図 17】



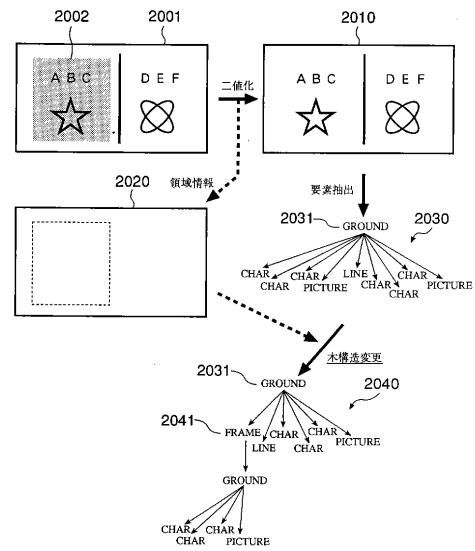
【図 18】



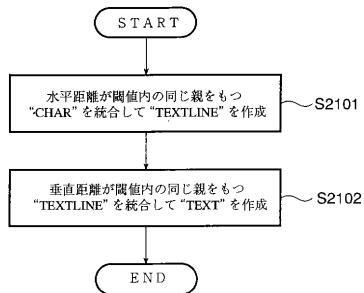
【図 19】



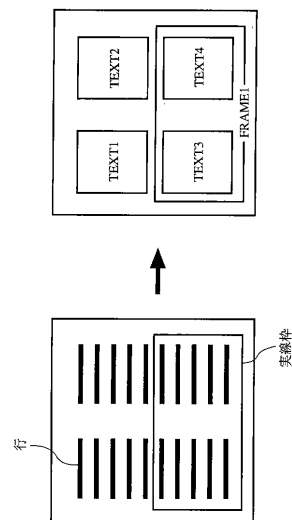
【図 20】



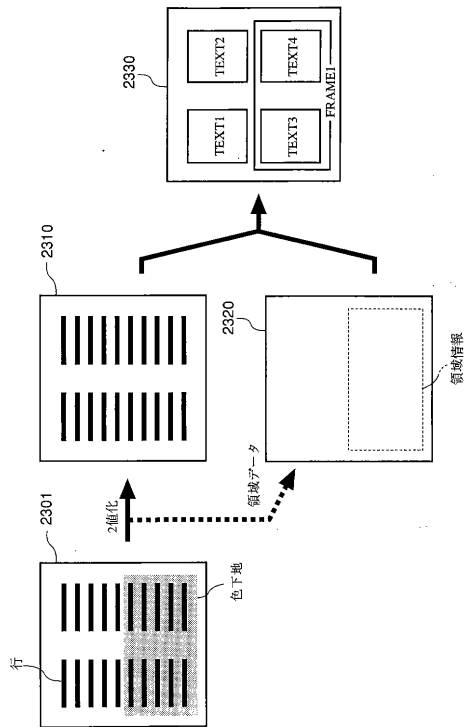
【図 21】



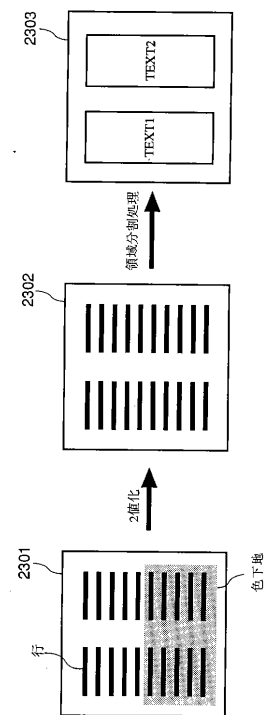
【図 22】



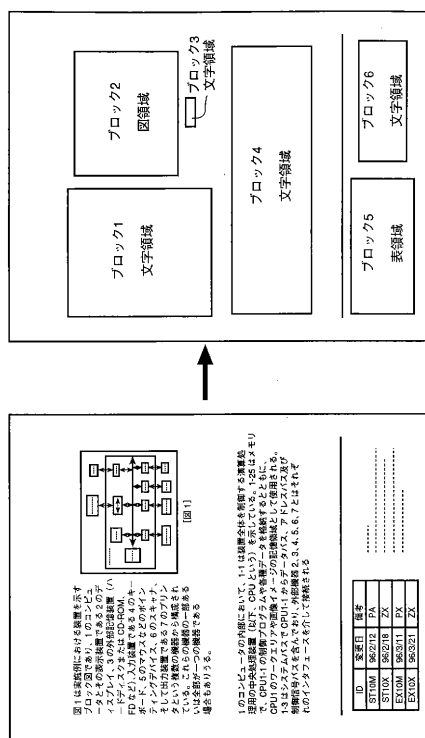
【 図 2 3 】



【圖 24】



【 図 2 5 】



ID	変更日	備考
ST10M	96/2/12	PA
ST10X	96/2/18	ZX
EX10M	96/3/11	PX
EX10X	96/3/21	ZX

フロントページの続き

(51)Int.Cl. F I
H 0 4 N 1/40 (2006.01) H 0 4 N 1/40 F

(72)発明者 金津 知俊
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

審査官 秦野 孝一郎

(56)参考文献 特開2000-132689(JP,A)
朴 栄碩 Young Seak PARK, マルチメディアにおけるパターンの認識・生成技術論文, 電子情報通信学会論文誌 (J75-D-II) 第2号 THE TRANSACTIONS OF THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 日本, 社団法人電子情報通信学会 THE INSTITUTE OF ELECTRONICS, INFORMATION AND COMMUNICATION ENGINEERS, 第J75-D-II巻